REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(LEAVE BLANK)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	29 April 1999	Professional Paper

4. TITLE AND SUBTITLE

A New Statistical Approach for Cyclic Life Tracking of Engine Critical Parts

5. FUNDING NUMBERS

6. AUTHOR(S)

Robert I. Kiang

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Air Warfare Center Aircraft Division
22347 Cedar Point Road, Unit #6
Patuxent River, Maryland 20670-1161

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Naval Air Systems Command
47123 Buse Road, Unit IPT
Patuxent River, Maryland 20670-1547

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This paper describes a flight data based statistical method that allows missing low cycle fatigue data in engine critical parts to be filled in with appropriate values that are not overly conservative but still ensure flight safety. The derived criterion resembles the Student's t-function in that it provides an estimate of the population mean given the sample mean. It differs from the Student's t-function in the following manner. This "t-function" is dependent upon two independent parameters, the number of samples and the percent of data captured; whereas, the Student's t-function depends only on the number of samples. When the capture rate is relatively low, this methodology can save many critical parts from early mandatory retirement.

14. SUBJECT TERMS

Statistical Cyclic life tracking Low cycle fatigue data t-function

15. NUMBER OF PAGES

14

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

A New Statistical Approach for
Cyclic Life Tracking of Engine Critical Parts

By

Robert L. Kiang
Naval Air Warfare Center Aircraft Division
Patuxent River, MD 20670

Abstract

This paper describes a flight data based statistical method that allows missing low cycle fatigue data in engine critical parts to be filled in with appropriate values that are not overly conservative but still ensure flight safety. The derived criterion resembles the Student's t-function in that it provides an estimate of the population mean given the sample mean. It differs from the Student's t-function in the following manner. This "t-function" is dependent upon two independent parameters, the number of samples and the percent of data captured, whereas the Student's t-function depends only on the number of samples. When the capture rate is relatively low, this methodology can save many critical parts from early mandatory retirement.

CLEARED FOR
OPEN PUBLICATION
29 apr 99
PUBLIC AFFAIRS OFFICE
NAVAL AIR SYSTEMS COMMAND
H. Howard

Introduction

A critical part of a gas turbine engine is defined as a component whose failure will cause an engine catastrophe with potential loss of the aircraft. Typical examples are fan hubs and turbine disks. Each critical part is assigned a maximum number of engine cycles, based upon design analysis and component testing, below which the part is guaranteed not to fail in low cycle fatigue (LCF). When a new engine goes into operation, the usage of each critical part is meticulously recorded. For flight safety, mandatory replacement is required when a critical part reaches its limit cycle.

Despite the wide spread use of automatic data recorders (ADR) in both commercial and military aircraft, 100% capture of the cycle counts are difficult to achieve. The missing hours are, essentially, the difference between the pilot reported flight hours and the data recorder hours. How to account for the corresponding missing cycles is the subject of this paper.

Using embedded algorithm and measured engine operating parameters (pressure, temperature, rpm, etc.), the data recorder calculates the cyclic usage of each of the critical part in real time. When the data are downloaded at a maintenance depot, an exchange rate, defined as the number of cycles used per engine operating hour, is then saved as permanent records. Variability in the exchange rate is relatively small in a commercial engine, but can be quite large in a military engine because of the necessary power variations in a particular mission. Accounting for the missing cycles amounts to assigning an appropriate exchange rate for the missing hours.

The task becomes a statistical exercise because one is interested in knowing the true mean value of the exchange rate had all the data been captured. The situation is akin to the Student's t-function from which one can ascertain the value of the population mean from the known sample mean and standard deviation. The t-function cannot be directly applied to the missing exchange rate problem, however, for two reasons. One is that the "population" referred to in the t-function is a sample of infinite size. The total number of recorded and missing exchange rates of a particular engine build in its operating life is far from infinite. Thus using the Student's t would give overly conservative results. The second reason, a more troublesome one, is that the Student's t-function has only one independent parameter, namely, the number of samples, n. In the missing exchange rate problem, there are two independent parameters: the number of known exchange rates in an engine build, n, and the capture rate, n/N, where N is the total exchange rate "population" had there been no missing data. The exercise, detailed in this paper, involves the generation of an equivalent t-function with two independent parameters using the available data from the data recorder. Use of this method can prevent premature retirement of expensive engine components while maintaining safe operation.

Current Fill-in Methodology

For the sake of flight safety, the missing exchange rate, or LCF stress cycles per flight hour, must be filled in with some values. Different engine manufacturers, for that matter, different engines may use slightly different fill-in methods. Some adopt the philosophy that any missing data are filled in with a worst case value, that is, the highest exchange rate ever recorded for that component. Others use a more statistically realistic approach by assigning a cut-off capture rate. The worst case value would still be used if the capture rate, over a certain period, is below this cut-off value. If the capture rate is above this value, then the average of all the captured exchange rates over this period is used as the fill-in value. The basis for this approach is of course the law of the average, namely, the average exchange rate value of the missing data, if they constitute a small fraction of the total data, is expected to be equal to the average of the captured data. The application of the capture rate cut-off method is illustrated below.

Figure 1 is a plot of the exchange rate data from a high pressure turbine disk of a specific turbofan engine versus time segregated by engine builds. Each point on the graph is an exchange rate calculated from the ADR downloaded raw data, namely, cycles used and accumulated flight hours over several flights. As noted on the top margin of this graph, this engine has gone through two builds in the 3-1/2 year period shown. Build 1 has also been installed in two different aircraft resulting in three data sets. In data set 1, there are 18 data points. But it is known from other flight records that 13 data points are missing, due to either ADR malfunction or other type of errors. This translates to a data capture rate of 58%. In set 2, no data was captured. In set 3, the capture rate was 62%. If one adopts a cut-off capture rate of 60%, then the 13 missing data points in set 1 are each assigned a fill-in exchange rate of 4.55 cycles per hour, a worst case value derived by some other statistical means. For data set 2, all data are filled in with a worst case value of 6.04. For data set 3, since the capture rate is above the cut-off, all missing data are filled in with an exchange rate of 3.95 cycles per hour which is the average of the captured data.

Filling in with the worst case values is probably over conservative (The cut-off method mitigates some of the conservatism by filling in with data average when the capture rate is high). To err on the conservative side is certainly commendable from the flight safety standpoint, but that means the parts will be retired prematurely resulting in increased maintenance cost of the engine. The methodology described below, modeled after the Student's t-law, provides a statistically based means to fill in the missing data without jeopardizing flight safety.

Student's t-law

Even though the explanation of the Student's t-law can be found in almost any statistical book, we summarize the essence of this law here for two reasons. One is that the statistical exercise to be presented in the following sections can be viewed as an extension of the Student's t-law. The second reason is that the t-law serves to introduce some of the terminology that will be used later.

For this application, in which the concern is how much the sample mean under-predicts the population mean, the Student's t-law can be written in the following form:

$$(\mu - X) / s > t_{\alpha,n-1} / (n)^{1/2} \qquad (1)$$

where μ = population mean
X = sample mean
s = sample standard deviation
n = number of samples
$t_{\alpha,n-1}$ = t-function that depends on α and n-1, the degree of freedom
α = probability in which the above inequality holds

With the tabulated t-function readily found in many of the statistical textbooks, the above formula allows one to calculate the difference between the population mean (which is unknown) and the sample mean as measured by units of standard deviation. A graphic representation of the Student's t-function is shown in Figure 2. If one chooses α to be 5%, the above formula says that there is a 5% probability that μ will exceed X by the amount dictated by the right hand side of Eq (1). This is equivalent to saying that one is 95% confident that μ will not exceed X by that amount. If one considers t as a function of confidence level (CL) instead of probability, then the Student's t-law can be recast into a form better suited for this discussion:

$$(\mu - X) / s < t_{CL,\,n-1} / (n)^{1/2} \quad (2)$$

Common sense tells us that when the number of samples is small, say 5 or 10, the sample mean, X, can be quite different from the true mean, i.e., the population mean, μ. As the number of samples increases, X would approach μ. This expectation is reflected in each of the curves in Figure 2, that is, the t value is large for small n, and reduces to a smaller, and eventually a near constant, for larger n. As an example, when n = 11 (or the degree of freedom equals 10), one has 95% confidence that μ - X is less than 0.54s (i.e., about half of the sample's standard deviation). If n = 101, the corresponding μ - X is less than 0.16s.

The assumptions underlying the t-law are (1) the samples follow a normal, i.e., Gaussian, distribution, and (2) the population is large (mathematically infinite). To test the normality of the data, exchange rates from several high pressure turbine disks have been plotted on Gaussian charts. One example of the charts is shown in Figure 3 (A perfect Gaussian distribution would show up as a straight line on this chart). The goodness of fit parameter n^2 of 0.933 indicates that the Gaussian assumption is acceptable. It has also been postulated that the exchange rate assumes a Weibull distribution. A Weibull chart is also included in Figure 3 as a comparison to the Gaussian chart. For the example shown, the data show that they are closer to a Weibull than a Gaussian distribution.

Example of Student's t Application

Over a period of 12 months, a specific engine was of one build and associated with one aircraft installation. There were 106 ADRS downloads during this period, each resulting in one exchange rate for each of the critical parts. For the HP turbine in this engine, the minimum exchange rate is 1.095 cycles/hour, the maximum 17.925, but the majority cluster around 5 and 6. We consider this set of exchange rates a data set with n = 106. It was also determined that the data capture rate during this period was 70%. This means that if the capture rate were 100%, there ought to be 151 data points, i.e., N = 151. According to the Student's t-law with 95% confidence,

$$(\mu - X) / s\,|_{106} < t_{CL,\,n-1} / (n)^{1/2} = 1.645 / (106)^{1/2} = 0.16$$

If, for some reason, one only had 50 ER values instead of 106, then the sample mean is expected to differ more from the population mean by:

$$(\mu - X) / s\,|_{50} < t_{CL,\,n-1} / (n)^{1/2} = 1.645 / (50)^{1/2} = 0.23$$

On the other hand, if one had all 151 values, then

$$(\mu - X) / s\,|_{151} < t_{CL,\,n-1} / (n)^{1/2} = 1.645 / (151)^{1/2} = 0.13$$

This example brings out a significant message conveyed by the Student's t-law. That is, the magnitude of (μ - X) is dictated by the number of data points. It is also noted that capture rate, the other parameter of interest, does not appear in the Student's t-law. The absence of capture rate from the Student's t-law is explained in the next section.

Finite versus Infinite Population Size

As mentioned earlier, the Student's t-law assumes the size of the population is infinite, i.e., $N \rightarrow \infty$. Since the capture rate is n/N, Student's t-law applies to the special case of capture rate being zero.

The number of exchange rate values in a particular data set will always be finite. If the capture rate of a particular data set is 100%, then the mean ER value would accurately reflect the LCF cycles of that engine during that period of operation. No fill-in ER is required. When the capture rate is, say, 90%, the corresponding mean ER is expected to be different from the "population" mean had 100% of the data been captured. The difference between the "sample" and the "population" mean is expected to grow as the capture rate becomes smaller. This is analogous to the Student's t-law except the independent variable is n/N instead of n. n, however, will remain a controlling parameter by itself for the following reason. For a specific n/N, the sample mean is expected to be closer to the population mean when n is larger (Intuitively, one would expect the average of two numbers taken from a population of 20 to be further away from the population average than for the case in which 200 numbers are taken from a population of 2000). What we need, therefore, is an alternative to the t-law that has two independent variables, n and n/N. Fortunately, the specific engine of interest has been in use for several years and there exist a large amount of field exchange rate data. We shall take advantage of the available data and generate a set of " t-functions" with two independent variables.

Methodology

The data available to us are contained in 282 data sets, each corresponding to either an engine build or an aircraft installation. That means for each of the four critical components of concern (LP compressor drum, HP compressor drum, HP turbine disk, and LP turbine disk) there are 282 sets of exchange rate values. Within each data set, there can be as few as two exchange rate values and as many as 130. The mean number of points in a set is about 40.

In this exercise, we assume the number of points in a data set is the full "population" and denote that number by N. Random removal of points from this set will generate many subsets each has n points and a corresponding capture rate (n/N). Note that even though N is far less than infinite, we still define the full set as the population for ease of identification. The corresponding average of all N points will still be denoted by μ. The average of a subset is again X. By the process of random removal, a (μ - X)/X versus n plot can be generated for a given N.

Figure 4 is such a plot for N = 80 (HPT data are used in this illustration). Within the 282 data sets, there are 36 sets that have 80 points or more. For sets that contain more than 80 points, excess points are removed by a random selection process so that each of the 36 sets starts with exactly 80 points, and a μ is calculated from these exchange rate values. Next, a point is removed, randomly, from each of these 36 sets. A sample mean, X, corresponding to n = 79, is then calculated, from which (μ - X)/X can be readily obtained. All 36 of these (μ -

X)/X values are plotted along the vertical line of n = 79. The process is repeated by randomly removing two points from each set resulting in further spreading of points along n = 78. As expected, the spread becomes wider and wider as n, or equivalently n/N, gets smaller. Assuming the vertical scatter seen along any value of n is Gaussian distributed, one can calculate the standard deviation in (μ - X)/X. Two curves corresponding to + and - twice that standard deviation are plotted in this diagram. Since 2-sigma encompasses 95% of the area under a bell shaped curve, this means 95% of the points should lie between the two curves. As far as exchange rate is concerned, one is more interested in the upper curve along which $\mu > X$. Since the Gaussian distribution is symmetrical about its mean, the probability that a point lies above the upper curve is only 2.5%. In other words, the upper curve marks the boundary at which one has 97.5% confidence that the sample mean X will not be greater than the population mean μ. As a matter of fact, it gives the expected amount in which X will be below μ. For instance, at N = 80 (that is the basis of Figure 4) and n = 40, or equivalently a capture rate of n/N = 50%, μ - X is expected (with 97.5% confidence) to be 0.1X. This upper curve, without the underlying points, is re-plotted in Figure 5.

Plots similar to Figure 5 have been generated for N = 90, 65, 50, 30, 20 and 10. From each of these 2-sigma plots, a threshold n, and its corresponding capture rate, for a specific value of (μ - X)/X can be determined. For example, in the case of N = 80 (see Figure 5), the threshold n and the capture rate are 18 and 0.23 respectively for (μ -X)/X = 0.2. Table 1 summarizes these threshold values for all N including this example of N = 80. For N = 50, three random reduction exercises were performed resulting in three slightly different two-standard deviation curves, hence three sets of values. These were done to get a measure of the scatter from different random reduction exercises.

Table 1

Threshold n and Capture Rate for (μ - X)/X = 0.2 with 97.5% Confidence

N	90	80	65	50	30	20	10
Threshold n	21	18	16	13, 16, 17	13	10	7
Threshold Cap. Rate	0.23	0.23	0.25	.26, .32, .34	0.43	0.50	0.70

Threshold n versus capture rate from Table 1 are plotted as diamonds, along with a linear regression, in Figure 6. For this linear fit, area to the upper right of the line marks the region where (μ - X)/X < 0.2, whereas that to lower left marks the region where (μ - X)/X > 0.2. Also shown in Figure 6 are points and lines for (μ - X)/X = 0.05, 0.10, 0.15, and 0.25.

Figure 6 is the desired "t-function" with two independent variables n and n/N. Application of this function is as follows.

1. Assemble an exchange rate data set for a specific engine build.
2. Determine the number of data points captured, n, and the number missing, N-n.
3. Calculate the average exchange rate X of the captured data.
4. Determine, from Figure 6, the value of (μ - X)/X corresponding to the n and n/N obtained in item 2. Interpolate if necessary.
5. Calculate the value of μ and use it to fill in all the missing data.

The linear regression lines in Figure 6 are made to converge at capture rate equals 1.0 and n = 0 because these values represent the physical limits of these two parameters. The lines' intercepts with the x-axis, corresponding to a capture rate of zero, are the equivalents of the Student's t-law. Indeed, Student's t predictions of (μ - X)/X from the individual 282 data sets (with a 97.5% confidence) are close to the results of Figure 6 at very small capture rates. (Note: A direct comparison to the Student's t prediction a n/N = 0 is not possible because for capture rates corresponding to n = 0, and for that matter 1, there is no scatter in the data and thus no standard deviation s. Neither the Student's t-law nor the "t-function" for two variables derived above is meaningful.)

Even though we have illustrated our methodology using only the HPT data, similar data analyses have been done for the HPC, LPT, and LPC data. Provided that sufficient exchange rate, or equivalently cyclic usage, data exist on any engine, the same methodology can be used to derive the "t-function" for a specific critical part.

A Critical Part Usage Example

As an example of how different fill-in methods would affect the cyclic usage of a critical part, we have chosen a low pressure turbine disk example. This disk has been installed in five different engine builds over its service life. The data capture rates by the automatic data recorder range from 0 to 91% as seen in Table 2. So fill in is needed in every one of the five builds. It has also been determined that the worst case fill-in value should be 11.9 cycles/hour. This value is used to calculate the worst case fill-in cycles based on the missing flight hours that needed to be filled in. The resulting total usage is then calculated to be 22,876 cycles. If on the other hand, one chooses to use a 60% capture rate as a cutoff above which mean exchange rate from the captured data is used as fill-ins (For builds with capture rates below 60%, the worst case fill-in value is still used), then the total usage is slightly less, at 22,768 cycles. The choice of a cutoff capture rate is rather subjective, but will influence the total cycle counts. Finally the "t-function" method is used to calculate the usage. As can be seen in Table 2, the worst case exchange rate is still used for build 3 because it has zero capture rate. The fill-in exchange rates for all other builds are determine by a n versus capture rate plot similar to that shown in Figure 6. Not only the subjectivity is removed, but in this case, the usage is reduced by 10% to 20,579 cycles.

Summary

A statistics based methodology is presented that determines of the proper fill-in exchange rate values when cyclic usage data are missing from the automatic data recorder. This allows a more accurate account of the total cyclic usage of a fracture critical part in a gas turbine engine thus promotes the overall flight safety.

Table 2 Fill-in Comparisons

Build No.	1	2	3	4	5	
No. of Dnload, n	13	6	0	99	97	
Pilot Hrs	116	81	337	896	828	
ADR Hrs	93	74	0	526	446	
ADR Cycles	779	592	0	4374	3815	
Cap. Rate	0.80	0.91	0.00	0.59	0.54	
Fill-in Hrs	23	7	337	370	382	
Av. ER, X	8.4	8.0	N/A	8.3	8.6	
Worst Case Fill-in						
Fill-in ER	11.9	11.9	11.9	11.9	11.9	
Fill-in Cycles	274	83	4010	4403	4546	
Total Cycles	1053	675	4010	8777	8361	$\Sigma = 22876$
60% CR Cutoff						
Fill-in ER	8.4	8.0	11.9	11.9	11.9	
Fill-in Cycles	193	56	4010	4403	4546	
Total Cycles	972	648	4010	8777	8361	$\Sigma = 22768$
"t-function" Method						
Fill-in ER	9.3	8.9	11.9	8.8	9.1	
Fill-in Cycles	214	62	4010	3256	3476	
Total Cycles	993	654	4010	7630	7291	$\Sigma = 20579$

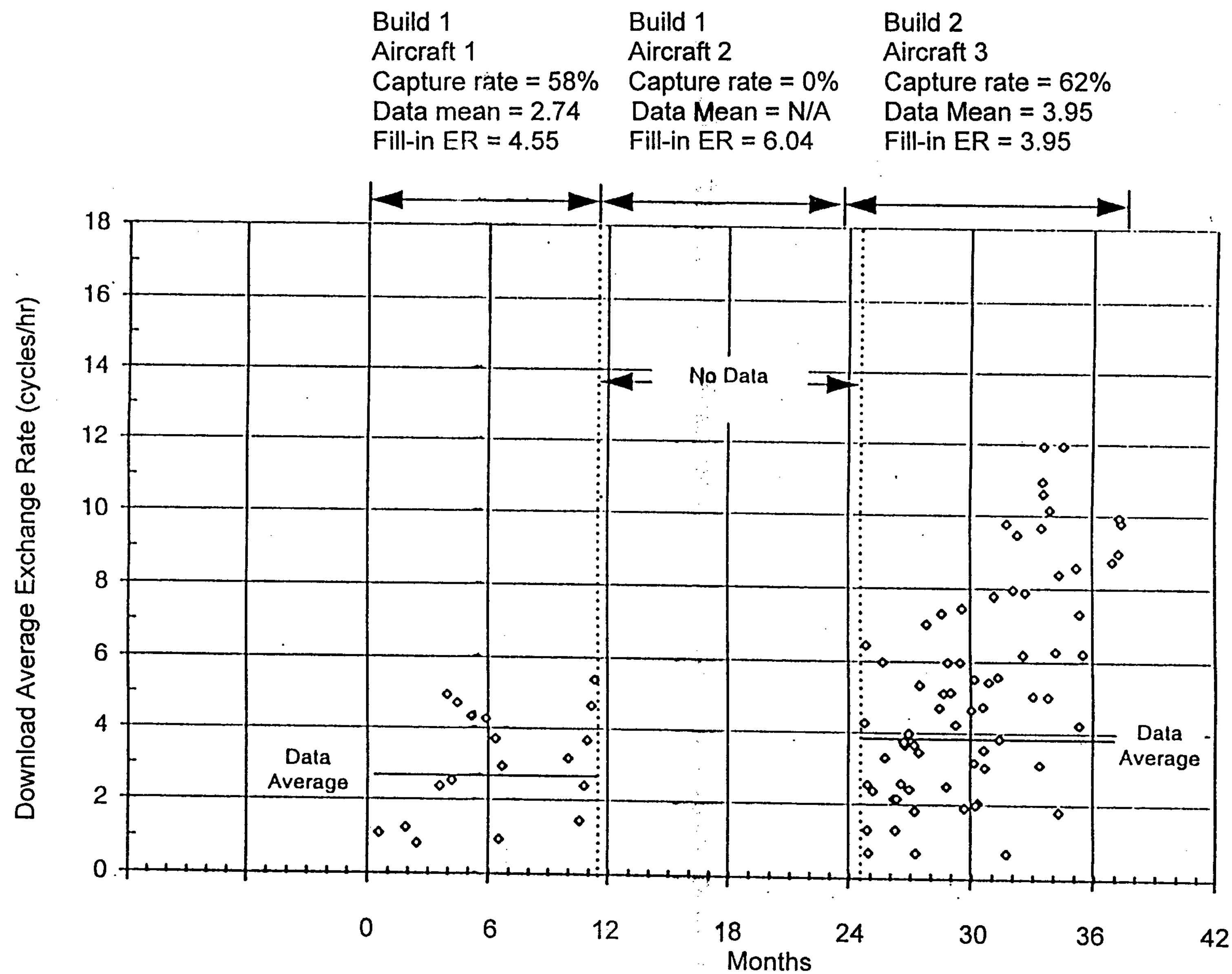

Figure 1 Exchange Rate Data of an HP Turbine Disk

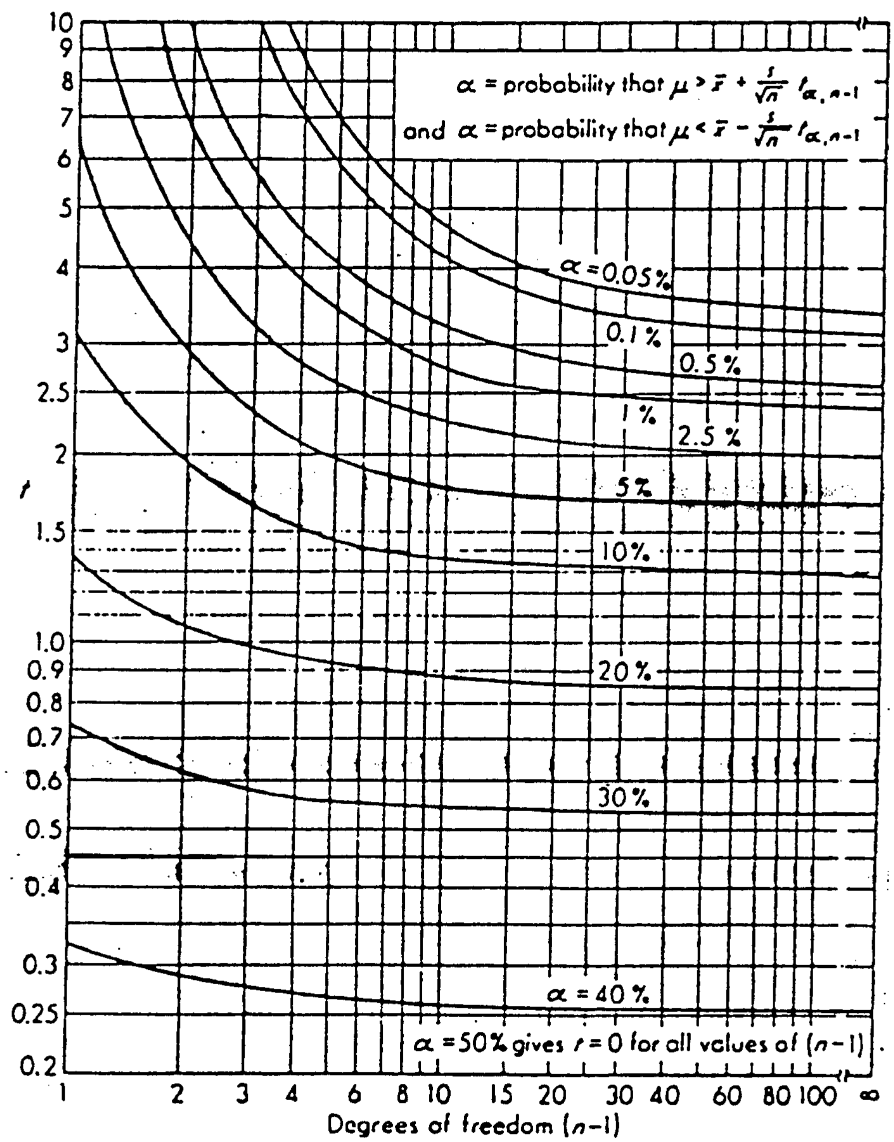

Figure 2 Student's t-function

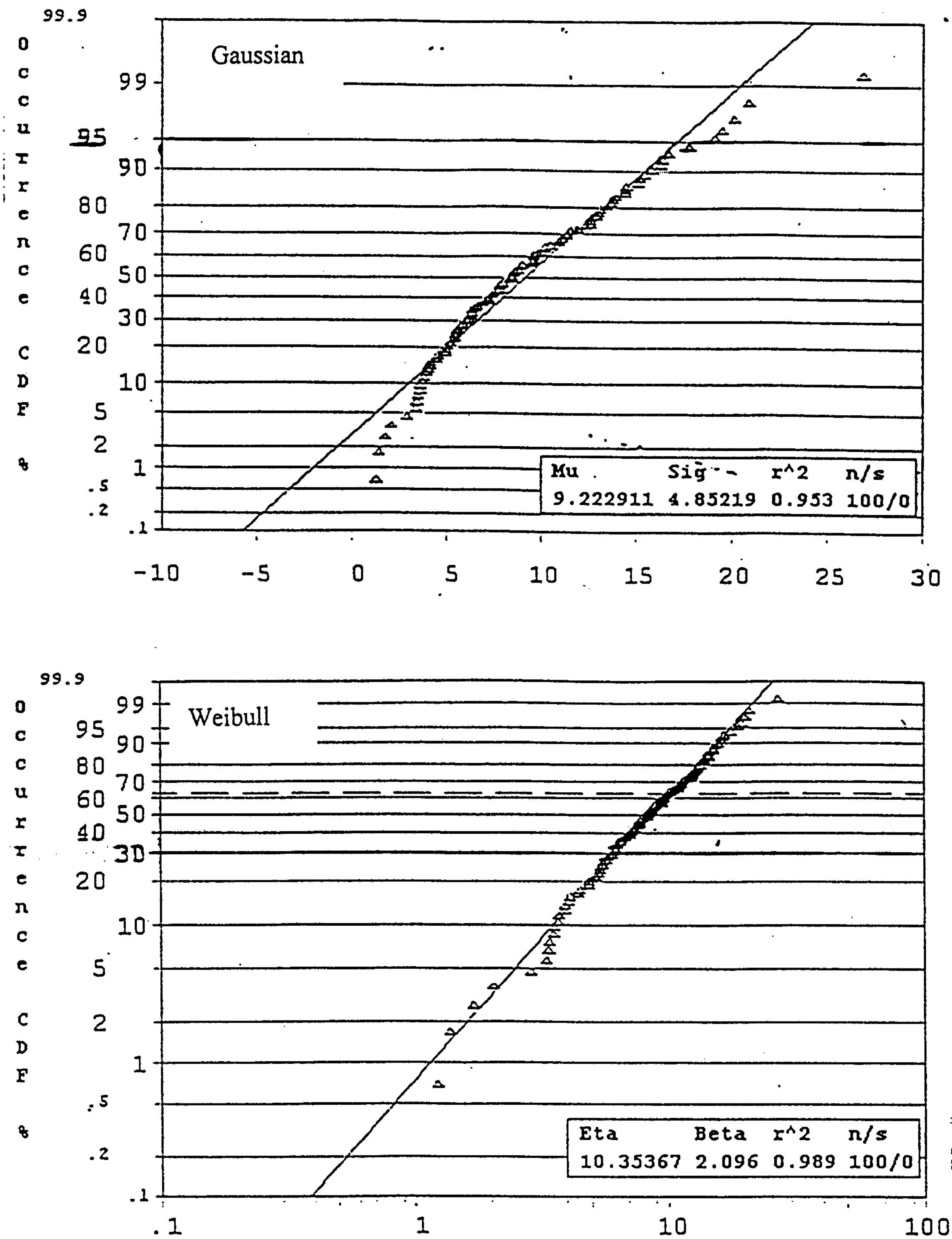

Figure 3 Exchange Rate Distribution – Gaussian vs. Weibull

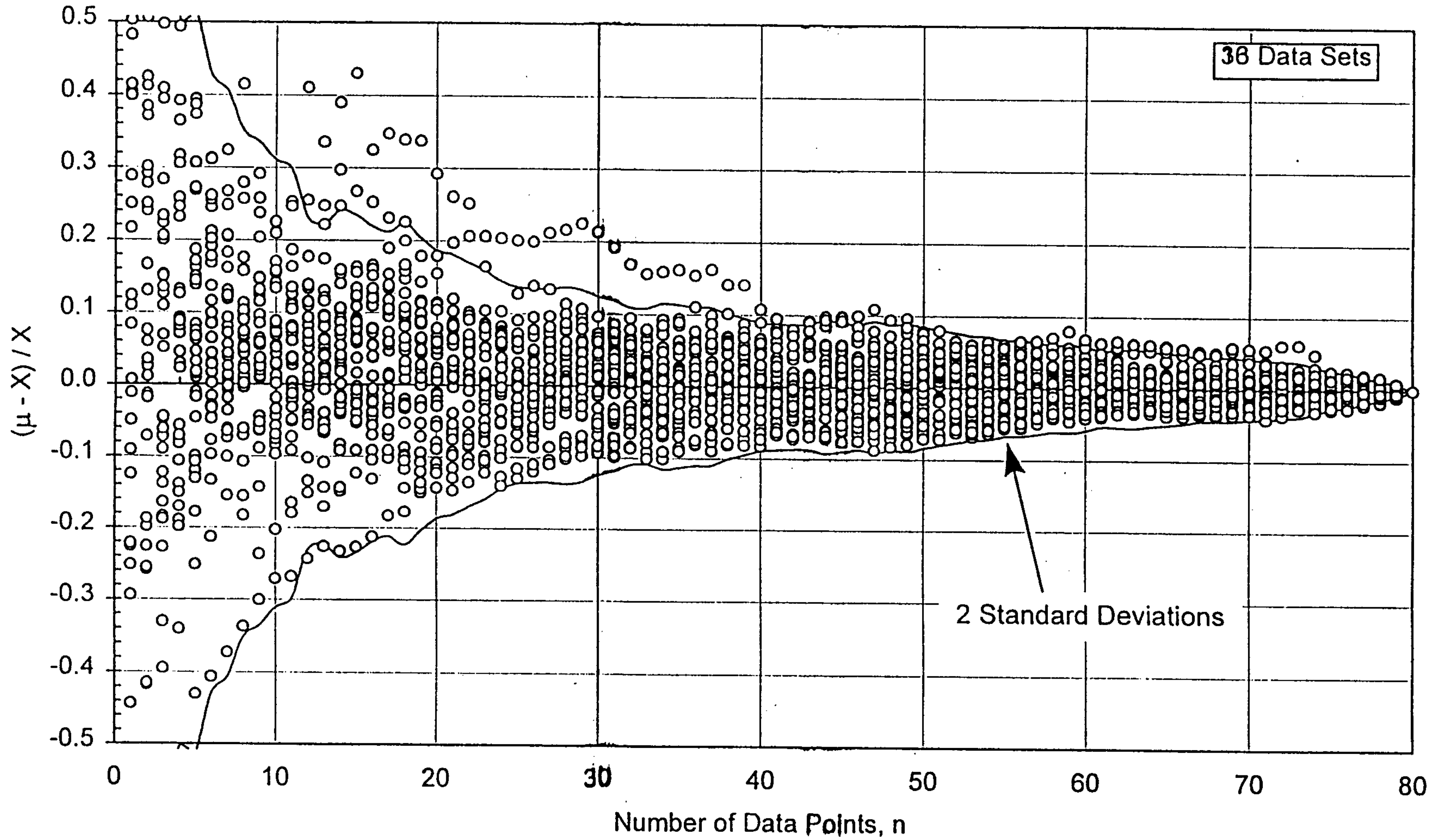

Figure 4

Effect of Capture Rate on Relative HPT Exchange Rate Means
(80 points per data set, i.e., N = 80)

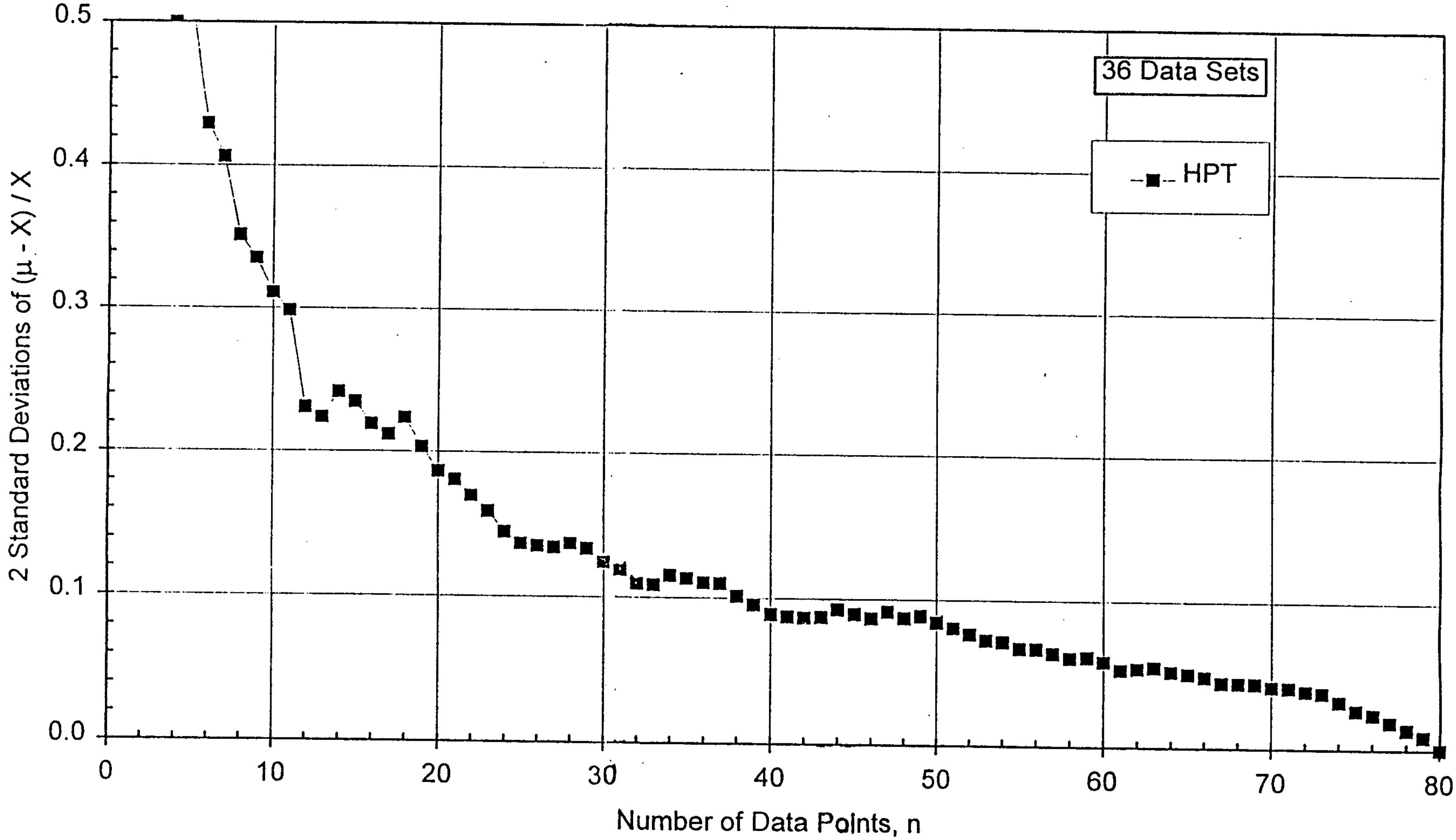

Figure 5

Two-sigma of the Relative Exchange Rate Means
(80 points per data set, i.e., N = 80)

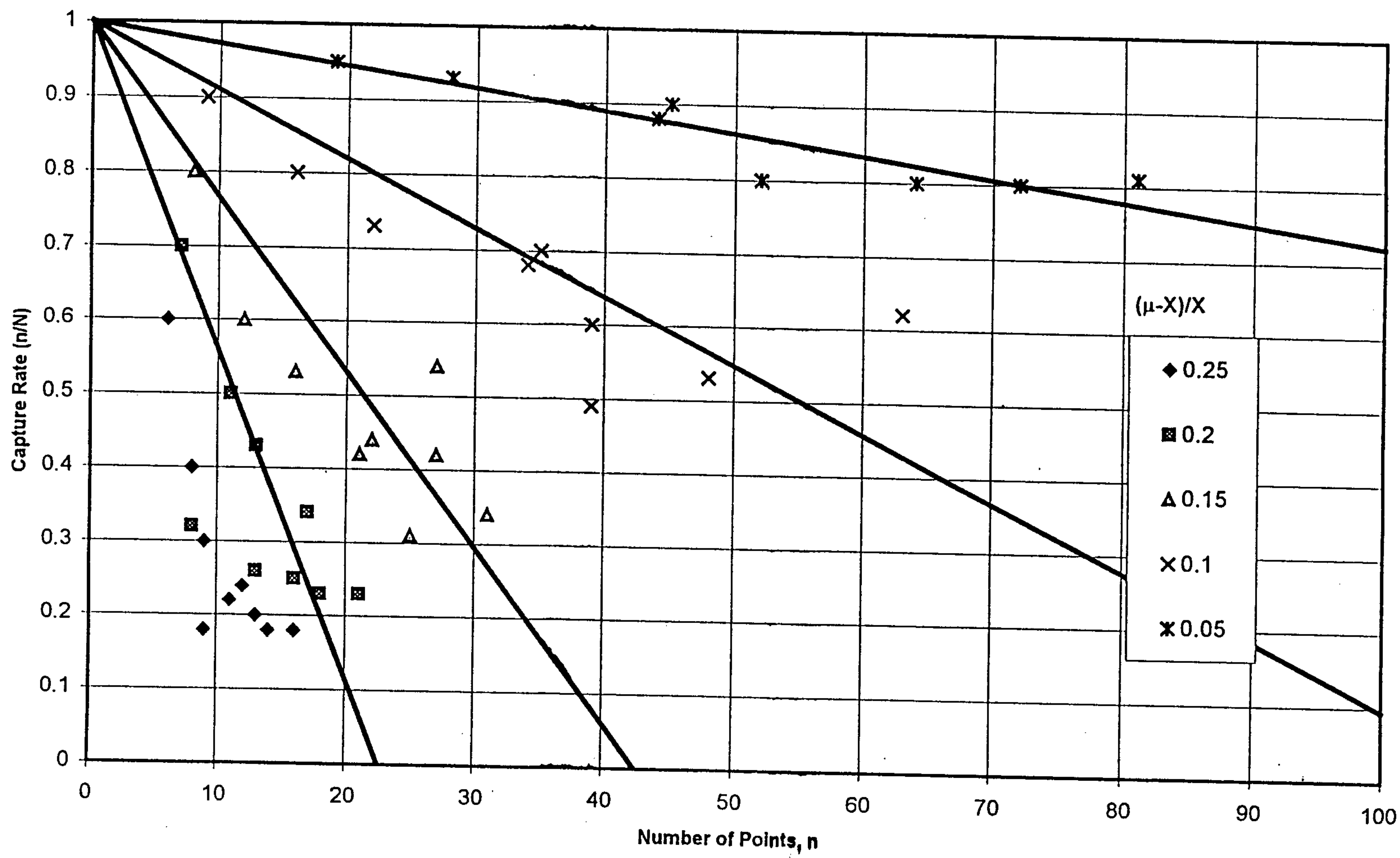

Figure 6 HPT Fill-in Exchange Rate Factor, $(\mu - X)/X$